A Bayesian framework for speech motor control

Jean-François Patri, Julien Diard, Pascal Perrier, Jean-Luc Schwartz

A BAYESIAN FRAMEWORK FOR SPEECH MOTOR CONTROL

Jean-François Patri[1], Julien Diard[2], Pascal Perrier[1], Jean-Luc Schwartz[1]

[1] GIPSA-lab, UMR 5216 - CNRS & Université Grenoble Alpes, France

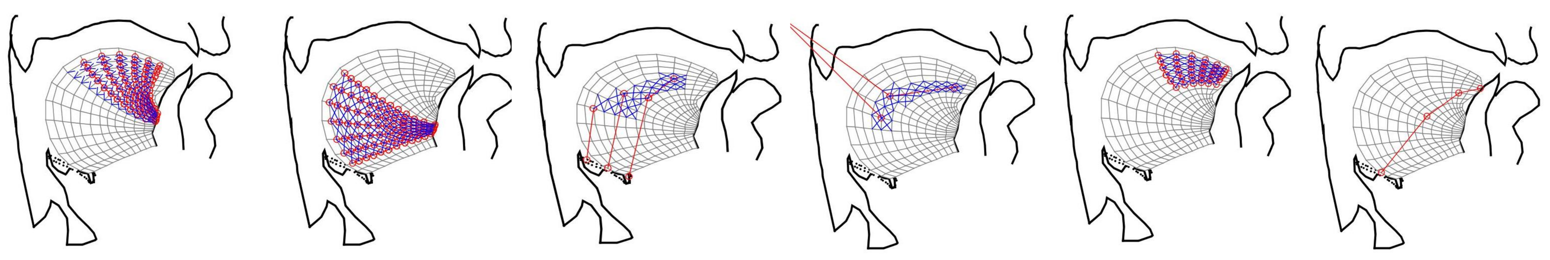
[2] LPNC, UMR 5105 - CNRS & Université Grenoble Alpes, France

Introduction

Speech is a skilled motor task achieving time series of goals within a timing that does not allow any online cortical processing of feedback signals. In addition, redundancy of the speech motor system makes the inference of motor commands an "ill-posed" inverse problem. Speech planning has been classically modeled within an optimal motor control framework by considering a feedforward control scheme coupled with a feedback controller. However optimal control schemes fail at accounting for token-to-token speech variability. In this context we proposed an alternative approach by formulating feedforward optimal control in a Bayesian modeling framework. We consider this approach to be appropriate for solving the ill-posed problem while accounting for the observed token-to-token variability in a principled way, and preserving the basic principles underlying the search for optimality without being explicitly driven by the minimization of a cost.

Context

Biomechanical model of the Tongue



Posterior Genioglossus | Anterior Genioglossus | Hyoglossus | Styloglossus | Verticalis | Inferior Longitudinalis

Control Scheme for a single phoneme

M → **S** → **Φ**

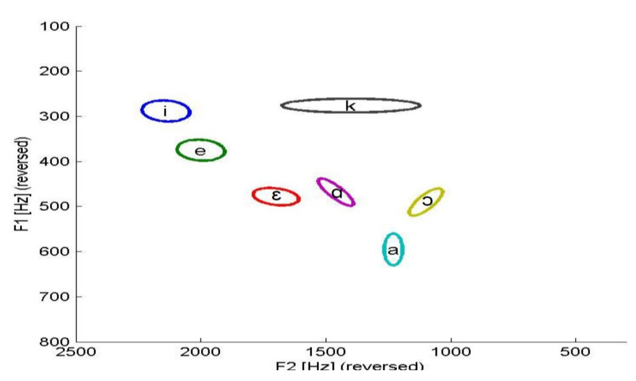
Motor-Control Variable
6 λ-parameters controlling activation of 6 muscles

Sensory variable
3 first Formants (peaks of spectral energy)

Phonemic Variable
Units of speech

Internal model
Radial Basis Function Network (RBF)

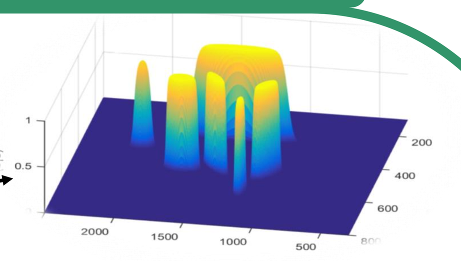
Target categorization
(Perceptual constraint)



Single Phoneme - Bayesian Model

Joint probability distribution

$$P(M\ S\ \Phi) = P(M)\ P(S \mid M)\ P(\Phi \mid S)$$

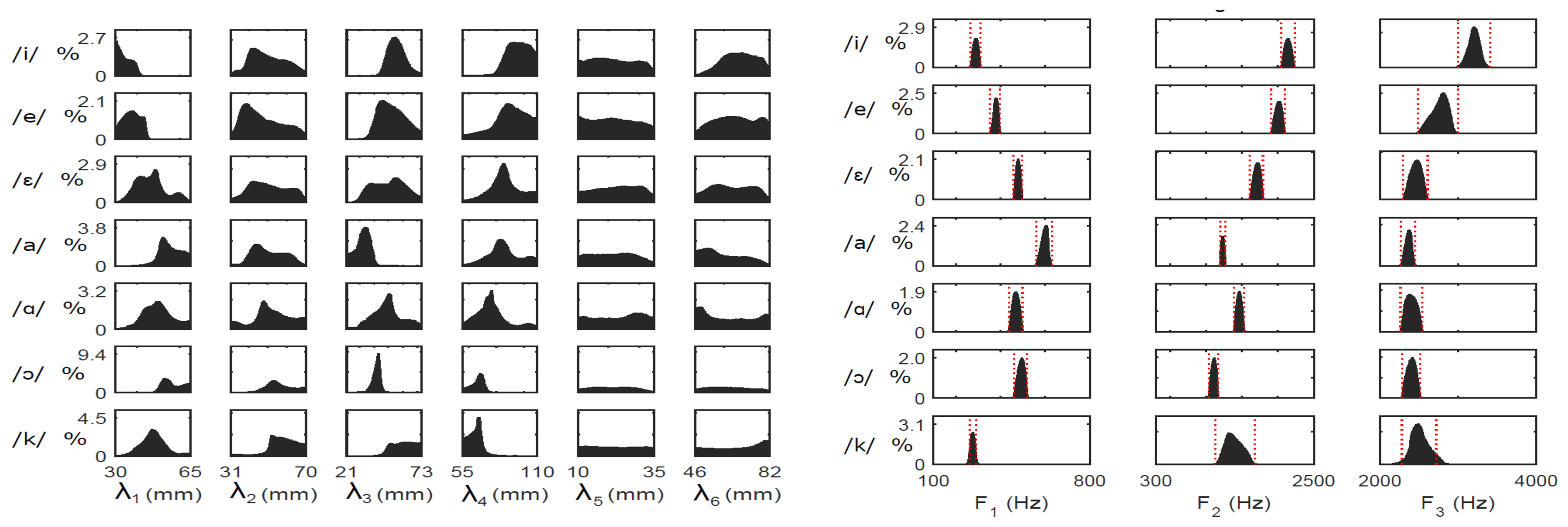


Uniform distribution | Dirac distributions on the RBF mapping | Bayesian inversion of Gaussian distributions $P(S|\Phi)$ corresponding to the ellipsoid target regions.

Inference of Control Variables

$$P(M \mid \Phi) \propto \sum_S P(S \mid M)\ P(\Phi \mid S) \propto P(\Phi \mid S^*(M))$$

Results



Distribution of inferred control variables for each phoneme → *Corresponding acoustic signal correctly distribute within the target regions.*

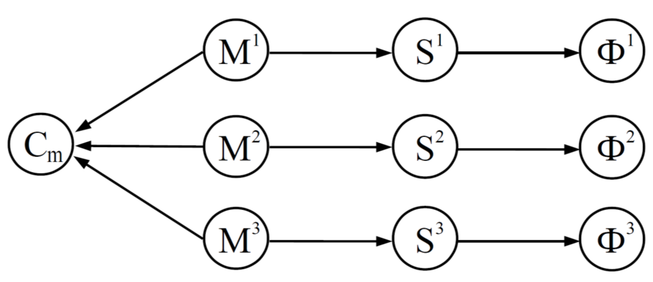
Sequence Planning - Bayesian Model

Assumptions

Control variables in sequence planning are selected in order to satisfy 2 constraints:

- **Perceptual constraint**: The corresponding acoustic output should correspond to the desired phonemic target
- **Motor constraint**: Laziness assumption: selected control variables should be as close as possible (for a 3 phoneme sequence: minimize the perimeter of the triangle that they define).

Joint probability distribution for a sequence of 3 phonemes



$$P(M^{1:3}\ S^{1:3}\ \Phi^{1:3}\ C_m) = P(M^1)\ P(S^1 \mid M^1)\ P(\Phi^1 \mid S^1)$$
$$P(M^2)\ P(S^2 \mid M^2)\ P(\Phi^2 \mid S^2)$$
$$P(M^3)\ P(S^3 \mid M^3)\ P(\Phi^3 \mid S^3)$$
$$P(C_m \mid M^3\ M^2\ M^1).$$

$\boldsymbol{C_m}$ is a binary variable that constrain control variables to be close when its state is ***L*** ("Lazy").

$$P([C_m = L] \mid M^3 M^2 M^1) = e^{-\kappa_M(|M^2-M^1|+|M^2-M^3|+|M^3-M^1|)}$$

$\boldsymbol{\kappa_M}$ is a parameter modulating the strength of the motor constraint

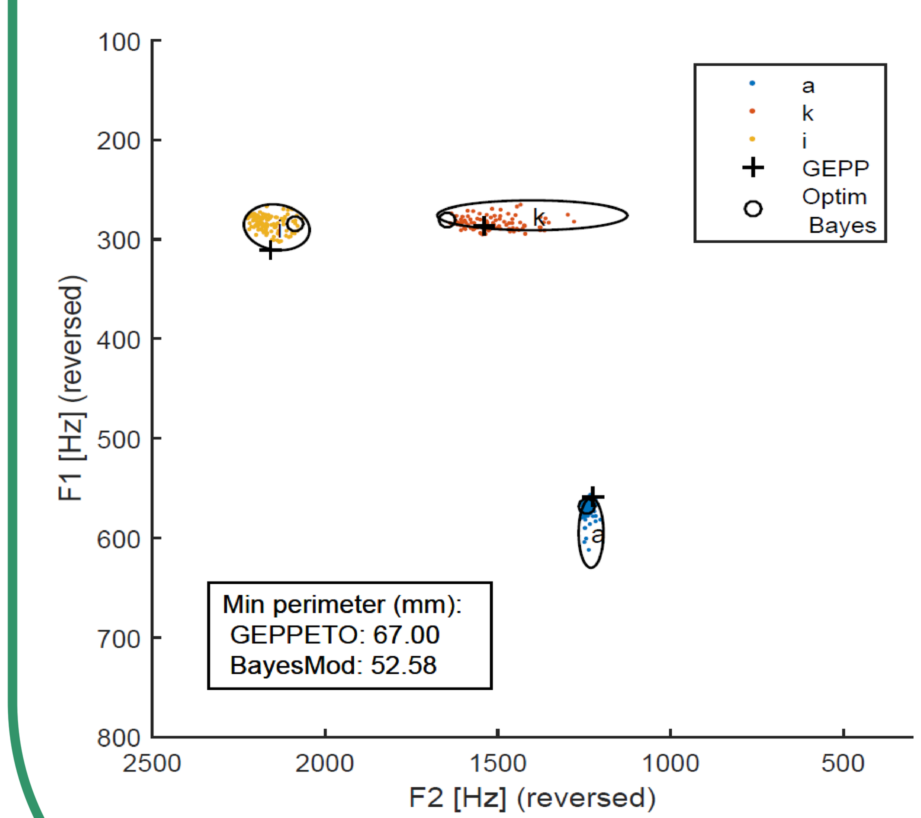
Inference of control variables

$$P(M^{1:3} \mid \Phi^{1:3}\ [C_m = L]) \propto P([C_m = L] \mid M^3 M^2 M^1) \prod_{i=1:3} P(\Phi^i \mid S^*(M^i))$$

Results

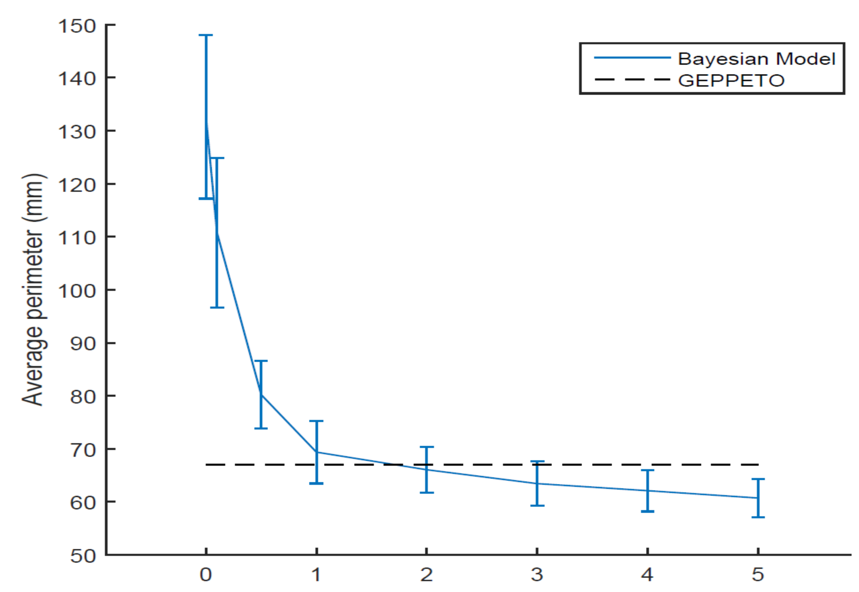
Comparison with an optimal control model

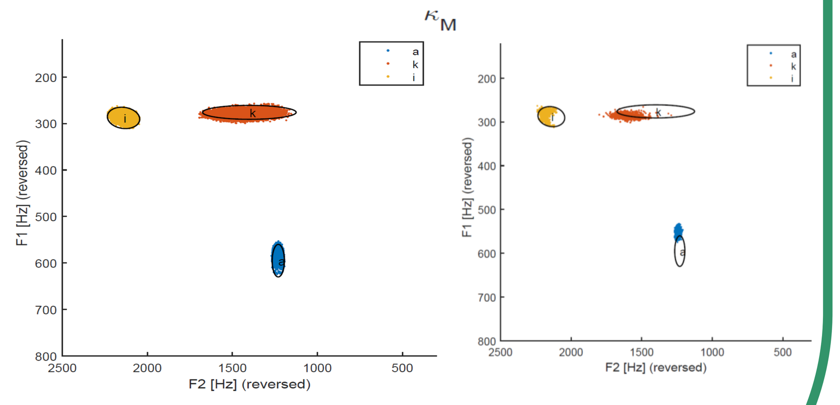
- GEPPETO is an Optimal Control Model that solves the same planning task under the same assumptions. It is based on the minimization of a cost function and leads to a unique solution
- The Bayesian Model leads to a distribution of solutions that are in agreement with the solution of GEPPETO.
- $\boldsymbol{\kappa_M}$ modulates the strength of the motor constraint. Relaxing the constraint leads to a decrease on coarticulation effect.



Min perimeter (mm):
GEPPETO: 67.00
BayesMod: 52.58

$\boldsymbol{\kappa_M = 1}$





$\boldsymbol{\kappa_M = 0,5}$ $\boldsymbol{\kappa_M = 5}$

Discussion

- **Equivalence of Models**
  - Both the Bayesian and optimal control models correctly infer control variables satisfying the constraints of the speech task.
    In addition, results are consistent with each other and indeed it can be shown that the Bayesian model includes GEPPETO as a special case.
- **Addressing redundancy and variability in formal terms**
  - The optimal control approach solves the redundancy problem with the specification of a unique and stereotyped solutions and leads to the elimination of all variability.
  - The Bayesian approach does not solve indeterminacy by suppressing all solutions but one, instead it characterizes every possible configuration by its probability to achieve the task. Redundancy is then solved by randomly selecting motor control variables under the corresponding probability distribution. The optimal achievement of the task is ensured in average.
  - Variability is an inherent consequence of the formalism. Furthermore, the variability generated with this approach has a specific structure that could be compared with experimental data.

The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007- 2013 Grant Agreement no. 339152, \Speech Unit(e)s", PI: Jean-Luc-Schwartz).